Chronic illness in a complex health economy: the perils and promises of downstream and upstream reforms

Jack Homer,[a]* Gary Hirsch[b] and Bobby Milstein[c]

Abstract

Chronic illness is the largest cause of death and source of health care costs in developed countries and a growing problem in developing countries. Here we build on past work in system dynamics and present a generic model of chronic illness, its treatment and prevention, applied to the U.S. population. The model explains the rising prevalence of illness and responses to it, including the treatment of complications and management activities designed to reduce complications. We show how progress in treatment and disease management has slowed since 1980 in the U.S., largely due to competition between health care payers and providers, resulting in price inflation and an unstable climate for health care investments. We demonstrate the impact of moving "upstream" by managing known risk factors to prevent illness onset, and moving even further upstream by addressing behaviors and living conditions linked to the initial development of these risk factors. Copyright © 2007 John Wiley & Sons, Ltd.

Syst. Dyn. Rev. **23**, 313–343, (2007)

Introduction

Chronic illness is a major health challenge facing all countries. It is the largest cause of death and source of health care costs in developed countries and has become a significant and growing problem in developing countries as well (Mackenbach, 1994; Olshansky and Ault, 1986; Wild *et al*., 2004; Mathers and Loncar, 2006). In the U.S., the Centers for Disease Control and Prevention (CDC) estimates that chronic illness is responsible for 70 percent of all deaths and 75 percent of all health care costs (CDC/NCCDPHP, 2007a). The aging of the U.S. population and increases in risk factors such as obesity suggest that chronic illnesses will be an even greater problem in future years. Already, according to the CDC, an estimated 32 percent of adults and 16% of children in the U.S. are obese (CDC/NCCDPHP, 2007b). The good news is that mortality rates from chronic illness have fallen significantly since 1970, dropping by about half for heart disease and stroke, for example (NIH/NHLBI, 2004). Even this good news must be tempered, however, since it means there are many more people living with chronic illness and its associated disabilities and health care costs.

Worldwide, the trends are even more stark. The World Health Organization (WHO) reports that 80 percent of deaths from chronic illness occur in

Jack Homer has been an independent System Dynamics consultant to private and public organizations since 1989. In the public arena, his work focuses on public health and health care policy. He was formerly an assistant professor at the University of Southern California and received a PhD in management from MIT. He also holds BS and MS degrees in applied mathematics and statistics from Stanford University, and received the Jay W. Forrester Award in 1997.

Gary Hirsch is an independent consultant with 35 years of experience in applying System Dynamics to complex problems in health care, education, social services and industry. He holds SB and SM degrees from MIT's Sloan School of Management in System Dynamics. His work in health care has focused on improved management of chronic illness, strategic planning for

---

[a] Homer Consulting, 3618 Avalon Court, Voorhees, NJ 08043, U.S.A. E-mail: jhomer@comcast.net
[b] Wayland, MA 01778, U.S.A. E-mail: gbhirsch@comcast.net
[c] Centers for Disease Control and Prevention, Atlanta, GA, 30341, U.S.A. E-mail: bmilstein@cdc.gov
* Correspondence to: Jack Homer.

Received February 2007; Accepted April 2007

System Dynamics Review Vol. 23, No. 2/3, (Summer/Fall 2007): 313–343
DOI: 10.1002/sdr.379
Copyright © 2007 John Wiley & Sons, Ltd.

lower- and middle-income countries (WHO, 2005). In many developing countries, deaths from chronic illnesses now outstrip mortality from traditional health concerns such as injuries and communicable diseases (Yach *et al.*, 2004); and rates of chronic illness are rising in the developing countries, creating an additional burden of disease on top of high rates of acute illness. The worldwide prevalence of diabetes, for example, is projected to rise from 171 million (2.8 percent) in 2000 to 366 million (6.5 percent) in 2030, with over 80 percent of the projected cases in 2030 occurring in the developing world (Wild *et al.*, 2004). Future economic development is expected to bring with it increased risk of morbidity and mortality tied to chronic illness and driven by growth in obesity, tobacco use, and other risk factors.

In most nations, health care systems are organized in a way that makes them hard-pressed to respond to chronic illness. The shortcomings of health systems in dealing with chronic illness include a failure to empower patients and involve them in their own care, a lack of linkages between the health care system and other community agencies that should be involved, misaligned incentives for providers, and a failure to invest in prevention (WHO, 2002). In the U.S., the Institute of Medicine has detailed changes needed in the health care system to cope effectively with the increasing burden of chronic illness, including consistent provision of evidence-based care, reorganization of clinical office practices to provide for longer visits needed for patient education and follow-up, attention to the needs of patients when seeking lifestyle and other behavioral change, and implementation of supportive information systems (IOM, 2001).

This paper begins with a review of past work in system dynamics (SD) concerning populations with chronic illness. It then presents a generic model of illness in a population and its treatment and prevention, applied to the U.S. population. This model encompasses not only chronic illness but all illness and injury, primarily because of data limitations discussed below. The distinction between chronic and acute conditions is a somewhat arbitrary one in any event. Some chronic illnesses can nowadays be cured quickly once they are discovered and, conversely, some acute infections or injuries, if not treated quickly, can become chronic problems. Combining all afflictions into a single model requires only that the rates of death and cure reflect the entire continuous distribution of illness, from the very short-lived to the very long-lived, and from the easily cured to the incurable. Although our model covers all manner of illness and injury, our focus is on those chronic illnesses which are long-lived and incurable, which are responsible for the great majority of health impairment in the U.S.

Our model explains the rising prevalence of illness as well as responses to it, responses which include the treatment of complications as well as disease management activities designed to slow the progression of illness and reduce the occurrence of future complications. The model shows how progress in complications treatment and disease management has slowed since 1980 in the U.S., largely due to a behavioral tug-of-war between health care payers and

health care delivery systems, and development of interactive learning environments for helping managers and providers better understand those systems.

Bobby Milstein, PhD, MPH, works at the Centers for Disease Control and Prevention where he leads the Syndemics Prevention Network and also coordinates planning and evaluation activities for emerging investigations and policy initiatives in chronic disease, environmental health, emergency preparedness, and health system improvement. With an academic background that combines cultural anthropology, behavioral science, and systems science, he guided the development of CDC's framework for program evaluation and often consults on the role of democratic processes in protecting the public's health.

providers that has resulted in price inflation and an unstable climate for health care investments. The model is also used to demonstrate the impact of moving "upstream" by managing known risk factors to prevent illness onset, and moving even further upstream by addressing adverse behaviors and living conditions linked to the development of these risk factors in the first place.

Applications of system dynamics to chronic illness

A number of applications of SD to chronic illness extending over three decades provide a foundation for the concepts discussed in this paper. Dental care and oral health was a focus of early work. The most expansive of these studies (Hirsch *et al.*, 1975) explored feedback relationships among the supply of personnel and the availability of care, the distinction between preventive and symptomatic care, the oral health status of a population and prevalence of dental disease, and the workload of dental practices. This study also analyzed the impacts of various dental manpower policies on oral health outcome measures including prevalence of decayed, missing, and filled teeth.

Several SD modeling efforts have focused on cardiovascular disease. Amodel developed for the State of Indiana Health Planning Agency (Hirsch and Myers, 1975) projected the prevalence of heart disease and stroke in the state and evaluated the potential impact of different programs for reducing the costs and mortality due to these diseases. The model represented multiple stages through which people move as they develop cardiovascular disease from predisposing conditions such as hypertension to undetected and nonacute illness, acute incidents such as heart attacks and strokes, and rehabilitation and recovery after such attacks. Simulations with the model illustrated the value of comprehensive programs that combine preventive interventions such as hypertension screening and treatment with improved acute care.

Another model of cardiovascular disease (Luginbuhl *et al.*, 1981) used a similar structure to examine the impact of investing more resources in prevention and rehabilitation rather than more elaborate technologies for treating acute myocardial infarction. The model demonstrated how prevention and rehabilitation could lower the costs of heart disease in the U.S. more effectively than new technologies that only marginally extend the lives of people who are in the later stages of the disease.

Diabetes is another area in which SD modeling has been used to study chronic illness in populations. A model developed for a community coalition in Whatcom County in the state of Washington (Homer *et al.*, 2004) portrayed patients flowing through several stages as they moved from being at-risk for diabetes into diabetes and its complications and moved from having their blood sugar levels not under control to under control. The model demonstrated how the right combination of interventions for prevention and treatment could reduce the burden of diabetes in terms of both mortality and cost. A similar population flow model of congestive heart failure—which, like diabetes, is another chronic illness producing high burden in the U.S.—was developed for Whatcom County and used for a similar analysis of interventions.

Another SD model of population flows in diabetes (Jones *et al*., 2006) was conducted for the CDC and developed with experts at the federal, state, and local levels in the U.S. This model is similar to the Whatcom County diabetes model in many ways, but enables a closer look at primary prevention by delineating the condition of moderately elevated blood sugar known as pre-diabetes and by portraying the significant influence of obesity (the leading modifiable risk factor for diabetes) on the onset rates for pre-diabetes and diabetes.

Other prominent SD models exploring the epidemiology of particular chronic conditions have addressed obesity (Homer *et al*., 2006) and smoking (Tengs *et al*., 2001).

Some SD modeling has considered chronic illnesses more generally, rather than focusing on a specific disease. A Health Care Microworld developed by the New England Health Care Assembly and Innovation Associates (Hirsch and Immediato, 1999) portrays a population at different ages as they develop and move through increasingly severe stages of chronic illness. Users of the Microworld can employ a variety of medical and non-medical interventions to influence these population flows, including interventions that can mitigate social, behavioral, and environmental risk factors for chronic illness.

The common feature of many of these earlier efforts is the focus on a population developing a specific illness and then moving through one or more stages of increasing severity and complications. Movement between these stages occurs at rates that depend on behavioral and environmental factors as well as demographic characteristics. The models allow for multiple points of intervention, both downstream after the disease process has ensued and upstream at points when disease incidence can still be prevented or the well-being of people better protected. A common lesson is the value of balanced strategies that include preventive programs as well as care and treatment to produce the most net benefit in both the short term and the long term. In the next section, we present a model that, aggregating across all illnesses, demonstrates the potential impacts of attempting to improve downstream care or upstream prevention and describes the economic mechanisms for such interventions.

A national-level model of downstream care and upstream prevention

Model scope and historical evidence

The shortcomings of the U.S. and other health care systems in dealing more effectively with chronic illness are systemic and not confined to particular localities or particular illnesses (IOM, 2001). They arise from the interactions of multiple stakeholders, including patients, providers, employers, third-party payers, makers of products, and regulatory and monitoring bodies as well as groups of ordinary citizens. Some of these actors are very influential, and their decisions can affect the health of an entire nation.

Accordingly, we have chosen to develop a model at the national level that aggregates across all illnesses to explore questions related to the evolution of downstream care and the potential benefits and costs of greater upstream effort. This model is based on data specific to the U.S., but its structure should be applicable to other countries. The past applications of SD to chronic illness have served as a useful background for this work, as have broader SD and non-SD studies that have considered the dynamics and economics of population health without a particular focus on chronic illness. The SD studies of this sort that have contributed to our thinking include (1) a simulation model of U.S. health care spending and finance (Ratanawijitrasin, 1993), (2) a simulation model of community health and the "syndemic" confluence of multiple interacting afflictions (Homer and Milstein, 2002, 2004), (3) a conceptual framework for thinking about U.S. health care reform (Hirsch *et al.*, 2005), and (4) a conceptual framework for thinking about upstream–downstream dynamics (Figure 3 in Homer and Hirsch, 2006). The influential non-SD studies of health include books by Starr (1982) and Heirich (1999) and articles by Weisbrod (1991) and Cutler *et al.* (2006).

The SD works considering the health system broadly have contributed useful ideas and hypotheses. In our present work, we have looked more closely at historical data and sought to develop a model capable of reproducing key elements of that history so that we may better understand its underlying causes. While such empirical grounding does not guarantee that the model is adequate and useful for exploring the future, it is an important step toward that end (Homer, 1996).

As we have gained familiarity with the historical data, we have come to focus our modeling effort on a perplexing question: Why, with the tremendous growth in health spending since 1960, is the health of Americans not better than it is? More specifically, why has the U.S. health care system, for all its size and capability, not managed to subdue chronic illness more effectively?

A key source of historical data has been the National Health Expenditure Accounts (NHEA) (CMS, 2007), which measure annual health spending in the U.S. by category. From 1960 to 2004, total health spending grew (in per capita, constant dollar terms) by a factor of eight, and as a fraction of gross domestic product (GDP) tripled from 5.2 percent to 16.0 percent. Note that we are no longer speaking of chronic illness alone: the NHEA data cover all health spending and do not distinguish between expenditures for chronic illness and those for acute illness and injury. Although estimates of national spending exist for some individual chronic illnesses such as diabetes (ADA, 2003), these are generally only on a one-time snapshot basis, and no comprehensive running audit of overall chronic illness spending is performed. Given this situation, and not wishing to abandon our desire to be empirically grounded, we have decided to expand the purview of the model to include all illness and injury, and not only chronic illness.

Total health spending grew rapidly from 1960 to 1990, slowed during the 1990s, then resumed more rapid growth in 2000. A consistent 82–85 percent of total health spending has been for what is known as personal health care (or what one might call health care consumption), which comprises hospital care (30 percent of health spending in 2004), nonhospital services (37 percent), drugs and health-related products and equipment purchased for individual use (13 percent), and miscellaneous personal health care (3 percent). Components of health spending in the NHEA other than personal health care include administration, public health, research, and capital investments. Rising costs for outpatient care have been responsible for much of the growth in health spending in the U.S. since 1980.

The recent historical record suggests the health of Americans has not improved as much as one might have expected from the dramatic growth in health care spending. We define illness or disease as a moderately or severely symptomatic biological or psychological condition—i.e., one associated with some reduction in perceived health-related quality of life. (A person with an asymptomatic or only mildly symptomatic condition is considered to be at risk for disease. Although not yet considered to have full-fledged disease, that person may be eligible for management or treatment of the risk condition.) Two of the CDC's large national annual health surveys—the National Health Interview Survey (NHIS) (CDC/NHIS, 2007) and the Behavioral Risk Factor Surveillance System (BRFSS) (CDC/BRFSS, 2007)—report the fractions of individuals describing their own health status as excellent, very good, good, fair, or poor. Research shows that these self-reported health metrics have desirable statistical properties and are predictive of adverse health events (Dominick *et al.*, 2002). The NHIS also publishes the self-reported prevalence of common chronic conditions. We have examined the reported results of other national surveys as well (Thorpe *et al.*, 2005; Hoffman *et al.*, 1996).

After considering all of these data sources, we have concluded that the NHIS sum of the poor, fair, and good health status categories (that is, people not reporting their health as excellent or very good) is the best indicator of the prevalence of illness as we have defined it, with a continuous span of reporting from 1982 to the present. Throughout this time period, this sum has remained within the relatively narrow range of 31–35 percent of the population, with some movement downward through 1990, upward through 1993, downward through 1998, then upward through 2004. Because the periods of downward movement are not consistent with some of the other measures described above, we are reluctant to emphasize the NHIS fluctuations before 1998. But the upward movement from 31 percent in 1998 to 33 percent in 2004 is clearly consistent with the other NHIS and BRFSS measures. We are thus confident in saying two things about the prevalence of illness since 1982: (1) it has not varied by much and certainly has not declined significantly, if at all; and (2) it has increased somewhat since the late 1990s.

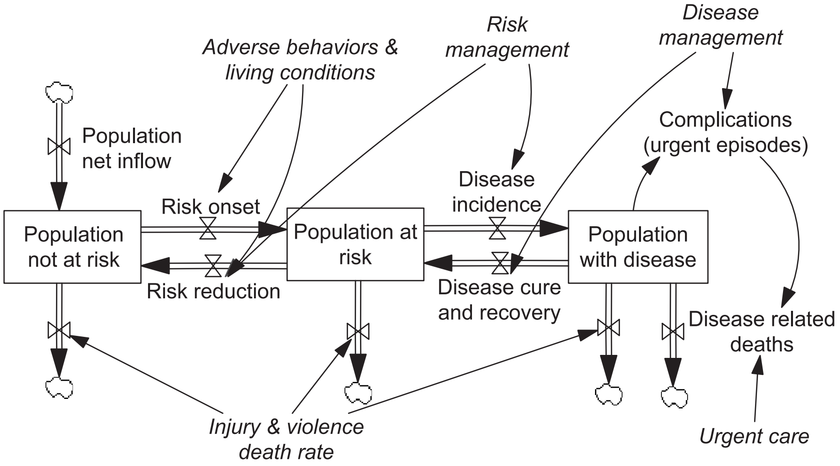
To address the question of why the U.S. has not been more successful in preventing and controlling chronic illness, we have constructed a simulation model that, although still a preliminary theory, can faithfully reproduce observed patterns of change in disease prevalence and mortality, and that can also reproduce the histories of the model's primary explanatory variables. The full model contains about 200 equations, including nine stocks and delay functions, 50 constants, and 11 exogenous time series. (The Vensim model is available in the online supplement at http://www.interscience.wiley.com/jpages/0883-7066/suppmat/sdr.379.html, or upon request from the authors.) Some of the exogenous time series ensure a closer model fit to history, while others represent potential policy levers. The exogenous time series do not affect the general findings discussed below; these findings are entirely determined by the model's feedback structure.

Conceptually, the model's hypothesized causal structure can be considered in three parts: (1) a population stock and flow structure; (2) feedback structure that explains the past and especially the growth of downstream care and spending; and (3) additional structure that can help explore the benefits and costs of upstream efforts to improve health.

Population stocks and flows

Figure 1 depicts all members of the population as being in one of three stocks: not at risk, at risk, or with disease. The population is increased by a net inflow rate, corresponding to births plus net immigration, and assumed for the sake of simplicity to flow entirely into the stock of population not at risk. The population is decreased by deaths, which are of two types: (1) deaths following

Fig. 1. Population stocks and flows as modeled


Risk management
Disease management
Adverse behaviors & living conditions
Population net inflow
Complications (urgent episodes)
Disease incidence
Risk onset
Population not at risk
Population at risk
Population with disease
Risk reduction
Disease cure and recovery
Disease related deaths
Injury & violence death rate
Urgent care

disease, affecting only the stock of population with disease; and (2) deaths from injury or violence, assumed to affect equally all three population stocks in the model. In 2003, the 1-year probability of death from injury or violence was 1 in 1743, and deaths from these causes accounted for only about 7 percent of all deaths (NSC, 2006; CDC/NVSR, 2006).

Disease prevalence is the fraction of the population with disease, while risk prevalence is the fraction of the population at risk. Disease is defined above. Risk refers to physical or psychological conditions or individual behaviors that may lead to disease. In particular, we have used as a proxy measure for risk prevalence the fraction of the adult population with one or more of the following cardiovascular risk factors: hypertension, high cholesterol, hyperglycemia, obesity, and smoking. BRFSS data from self-reports indicate that this measure of risk prevalence grew continuously during the period 1991–1999, rising from 58 percent to 62 percent (Greenlund *et al.*, 2004).

Flows between the stocks, as well as disease-related deaths, may be affected by certain actions and factors that we will discuss in the remainder of this paper. The disease-related death rate is affected by the effectiveness of *urgent care* for disease complications, generally involving hospitalization. The frequency of complications, in turn, is affected by the effectiveness of *disease management*. In some cases, effective disease management may increase the likelihood of disease cure or recovery; this is certainly true for many acute infectious diseases and can also be true for chronic diseases, as in the case of organ transplantation or cancer chemotherapy. Effective *risk management* can reduce the flow of people from risk to disease, and may also in some cases allow people to return to a condition of being no longer at risk. Such management may include changes in nutrition or physical activity, stress management, or the use of medication.

Flows of risk onset and risk reduction are affected by *adverse behaviors and living conditions*. Adverse behaviors may include poor diet, lack of physical activity, or substance abuse. Adverse living conditions can encompass many factors, including crime, lack of access to healthy foods, inadequate regulation of smoking, weak social networks, substandard housing, poverty, or poor educational opportunities. In calibrating the model, we have found that the rise in risk prevalence for 1991–1999 described above can be explained by assuming that the onset of risk due to adverse behaviors and living conditions increased by 30 percent from 1980 to 1995, and by another 5 percent through 2005. The timing and shape of this increase correspond well to the apparent historical pattern of growth in net caloric intake that has driven the rise in obesity in the U.S. since the late 1970s (Homer *et al.*, 2006).

Downstream loops

Figure 2 presents a theory of the growth of downstream care and spending. This growth is affected by changes in disease prevalence, as well as by changes

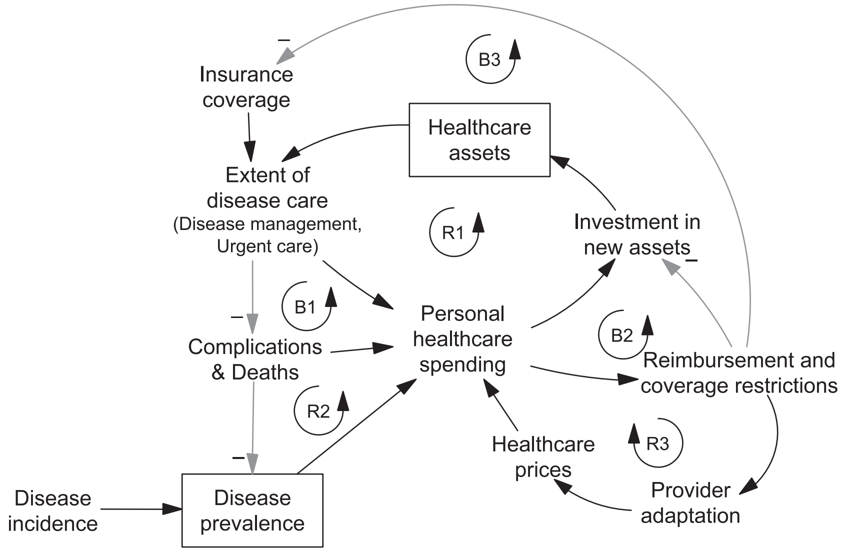

–
Insurance coverage
B3
Healthcare assets
Extent of disease care (Disease management, Urgent care)
R1
Investment in new assets
–
–
B1
Personal healthcare spending
B2
Complications & Deaths
Reimbursement and coverage restrictions
R2
Healthcare prices
R3
–
Disease incidence
Disease prevalence
Provider adaptation

Fig. 2. Theory of downstream health care system expansion and adaptation

Key to downstream feedback loops:

R1: Health care revenues are reinvested for further growth
B1: Disease management reduces need for urgent care
R2: Disease care prolongs life and further increases need for care
B2: Reimbursement restriction limits spending growth
B3: Insurance denial limits spending growth
R3: Providers circumvent reimbursement restrictions

in the extent of care (disease management and urgent care) and in health care prices.

What are the drivers of extent of care? For the purposes of our model, we have reduced a large literature on health care quality (see IOM, 2001), of which extent of care is a part, down to just two factors: the abundance of health care assets, and insurance coverage. By health care assets we mean the structures and fixed equipment used directly for health care or for the production of health care products, as well as the human capital of personnel involved. A greater abundance of assets nationwide means that a larger number of people have access to a broad array of medical services, but beyond a certain point some of that greater abundance represents duplication, and as a result one reaches a point of diminishing returns to extent of care.

By insurance coverage we mean the fraction of the population with some form of health care insurance, either with a private insurer or through a government plan. (Government plans are available in the U.S. for those with lower income, the elderly, the disabled, and for military personnel and war veterans.) The uninsured are less likely than the insured to receive health care services. The effect of insurance on extent of care is modeled as being relatively strong in the case of disease management services, for which the vast majority of providers require payment (something most of the uninsured cannot afford); while the effect is weaker in the case of urgent-care services, reflecting the fact that hospitals in the U.S. are required to provide emergency department access even to patients unable to pay.

Health care assets The model includes two separate stocks of health care assets which differ in terms of their uses: a stock used for disease or risk management, and a stock used for urgent (complications) care. These two stocks have likely grown at different rates at different times over the years. Distinguishing these two stocks and their different growth rates in the model has helped us to explain the evolution of health care spending evident in NHEA data, from more rapid growth in urgent care in the 1960s and 1970s, focused on hospital-based life-saving interventions, to more rapid growth in disease and risk management since about 1983, focused on the development and use of diagnostic equipment and pharmaceuticals.

To calibrate the asset sector of the model, we have looked primarily to NHEA data on investments in structures and equipment (S&E). (We have found no dataset on human capital in health care that is complete and can be harmonized with the data on structures and equipment.) In particular, we have estimated (via spreadsheet calculations) the net value of health care assets by accumulating the S&E investments over time, decrementing for obsolescence or depreciation at an assumed rate of 5 percent per year, and initializing in 1960 at a level that permits smooth early growth of the estimate. The resulting estimate grows at an average rate of 4.1 percent per year during 1960–1980, and 3.2 percent per year during 1980–2004. This growth in assets is consistently less than that of personal health care spending (consumption), which grew 5.4 percent per year during 1960–1980 and 4.3 percent per year during 1980–2004. We hypothesize that this difference reflects a decline over time in the fraction of health care revenues reinvested in assets. In fact, we have found it is possible to reproduce the estimated time series for health care assets by assuming that the revenue reinvestment rate declined from 13–14 percent in the 1960s to 10 percent in 1980 and to 6 percent in 2004.

Why should the revenue reinvestment rate have declined in this way? We suggest that the cutback in investment has been the response by potential investors to various forms of cost control, including the restriction of insurance reimbursements, which affect the providers of health care goods and services.[1] With increasing controls and restrictions, these potential investors face greater risk and uncertainty about the future return on their investments, and the result is a greater reluctance to build a new hospital wing, or to purchase an expensive new piece of equipment, or even, at an individual level, to devote a decade or more of one's life to the hardship of medical education and training.

Taking one step back, why is it that cost controls started to take hold in the 1970s and not earlier? Several authors (e.g., Starr, 1982; Eckholm, 1993; Heirich, 1999) have described how economic power, starting in the 1970s, shifted from providers of medical care, who had been allowed to act freely for many decades, to employers and public agencies desiring to rein in costs. As Paul Starr (1982) puts it, "Until the 1970s . . . practitioners, hospitals, researchers, and medical schools enjoyed a broad grant of authority to run their own affairs. In the 1970s the mandate ran out." Max Heirich (1999) describes this shift as a reaction to the growth in health care costs relative to the rest of the economy beginning in the 1960s: "Where for decades [the costs of American health care] had consumed between 3.5 and 4.5 percent of GNP . . . by 1960 its share of the GNP had increased to 5.3 percent . . . and health care's share of GNP increased to 7.3 percent in 1970 . . . The American health-care system's non-equilibrium growth in costs now affected the rest of the economy."

Health insurance coverage While some employers have reacted to high health care costs by selecting less generous, more cost-restrictive insurance plans for their employees, others have taken the more drastic action of not providing coverage at all to many of their workers. Surveys of insurance coverage taken annually since 1987 (U.S. Census Bureau, 2007) show that the fraction of the U.S. population covered to some degree by private (employer-provided or self-purchased) insurance fell from 75.5 percent in the late 1980s to 70 percent through most of the 1990s, rose briefly to 72 percent during 1999–2000, then declined again to 67.7 percent by 2005. This decline in private coverage is a serious matter affecting the ability of tens of millions of Americans to gain access to regular, good-quality health care. However, there is another dimension to the insurance story, and that is the growth of government-provided insurance. This growth started with the passage in 1965 of the federal Medicare and Medicaid programs to provide coverage for elderly and lower-income people, respectively. The Medicaid program in particular has grown over the years in terms of the fraction of the population it covers, from about 8.4 percent in 1987 to 12 percent in the early 1990s, declining to 10 percent in the late 1990s, and then rebounding to 13.0 percent by 2005. Thus, the Medicaid curve has for nearly 20 years moved consistently in a direction opposite to the curve for private insurance: a decline of 7.8 percentage points in private coverage has been countered by an increase of 4.6 percentage points in Medicaid coverage. As a result, the fraction of the population with any insurance coverage, private or public, has fallen by only 3 percentage points, from 87.1 percent in 1987 to 84.1 percent in 2005. Clearly, many of the people who have lost coverage from their employer or as a result of changing jobs, primarily wage-earners in lower paid positions, have been able to switch over to Medicaid as a fallback.[2]

Let us review the story of the health care system's evolution told thus far by walking through the hypothesized feedback loops in Figure 2:

- Loop R1 shows how the funds generated by health care lead to more investment in assets, and how the application of these new assets in the form of more extensive care can generate even more funds to support further growth. Even in today's more restrictive climate, this loop remains central to the story of progress in health care.
- Loops B1 and R2 show how more extensive care has effects on health and longevity that can moderate or reinforce Loop R1. Loop B1 indicates that increased disease management can prevent costly complications and thereby *reduce* spending and the need for investment in new assets for urgent care. Loop R2 indicates, however, that insofar as more extensive care prolongs life for people with disease, it tends to increase disease prevalence and thereby *increase* spending and investment in health care assets.
- Loop B2 shows how rising personal health care spending as a fraction of GDP triggers a backlash from employers and other payers, resulting in a more restrictive reimbursement climate that can suppress the rate of investment in new assets and thereby slow the growth in health care costs, although at the same time slowing further growth in the extent of care.
- Loop B3 shows how the denial of insurance coverage by some employers in reaction to high health care costs appears to be another route for slowing the growth in those costs, although, like Loop B2, it also slows growth in the extent of care.[3]

Taken together, one may view these loops as the story of a health care system that favors growth and investment until the resulting costs get to a point where further increases are perceived to be no longer worth the expected incremental improvements in health and productivity. That does not by itself sound like a story of dysfunction but rather one of progress followed by goal-seeking behavior. There is a potential for dysfunction in Loop B3, where a reduction in insurance coverage can drive up the unreimbursed costs of hospitals (resulting in a burden on the general public), and also create a situation of health inequity that separates the uninsured poor from the rest of society. But, although the insurance gap is certainly a matter of concern, that gap has been with us for decades, and its growth by 3 percentage points since 1987 is not by itself alarming. Because of this small magnitude of change, declining insurance coverage is unlikely to contribute much toward answering our question of how it is that health care spending can keep growing without doing much to improve health for the majority of the population.

HEALTH CARE PRICES To find a more compelling causal mechanism behind this sort of system failure, we must go one step further and consider the dynamics of health care prices. Medical care is one of eight major groups in the Consumer Price Index (CPI) computed by the U.S. Bureau of Labor Statistics (BLS), measuring retail price changes over time "for a constant quality, constant quantity market basket of goods and services" (BLS, 2007). The medical care

CPI combines four major components, with approximate importance weights for 2005 as follows: professional services (2.8), hospital services (1.6), drugs and other personal use products (1.5), and health insurance (0.4). The medical care CPI has grown more rapidly than the general CPI for the overall economy, especially since 1980. For 1960–1980, inflation in medical care prices averaged 6.2 percent compared with general inflation of 5.3 percent, while for 1980–2004 inflation in medical care prices averaged 6.1 percent versus general inflation of 3.5 percent. Consequently, a fixed market basket of medical care goods and services costing $100 in 1960 had risen to $1391 in 2004, while a market basket for the general economy costing $100 in 1960 had risen to $638 in 2004.

Why has health care inflation exceeded that of the general economy? We have considered various possible explanations for why costs should have gone up so rapidly, particularly since 1980, for a given quality of care. These include increasing costs for drug development; more gadgetry in medical technology; the increased practice of “defensive medicine” by providers to avoid lawsuits alleging malpractice; the increase in medical malpractice insurance premiums; the shift of many procedures from inpatient settings to outpatient settings where prices may be less tightly regulated; and the use by providers of various methods to maintain their incomes in the face of greater restrictions on reimbursement. Although all of these phenomena have contributed to health care inflation, not all have contributed with sufficient magnitude or with the timing necessary to explain the historical pattern. One phenomenon that does appear to have such explanatory power, and which we have centered on for the purposes of this study, is the last one listed above, described in Figure 2 as “provider adaptation”, or elsewhere as “the target income hypothesis” (Ratanawijitrasin, 1993, p. 77) or “the behavioral response” (Peter Passell in Eckholm, 1993, p. 285).

A variety of studies since the late 1970s provide strong support for the idea that, in response to cost containment efforts, providers may “increase fees, prescribe more services, prescribe more complex services (or simply bill for them), order more follow-up visits, or do a combination of these” (Ratanawijitrasin, 1993). Specific billing practices that can circumvent cost containment efforts include “upcoding” (billing with procedure codes that receive higher reimbursement rates) and “unbundling” (billing a single procedure in multiple parts to achieve a higher total) (Eckholm, 1993). Many tests and procedures are performed that contribute little or no diagnostic or therapeutic value, thereby inflating the cost per quality of care delivered. Writing in the *New York Times* in April of 1989, the former Secretary of Health, Education, and Welfare, Joseph Califano, Jr., claimed that “Americans would spend about $155 billion in 1989 for tests and treatments that would have little or no impact on the patients involved” (Heirich, 1999, p. 97). If correct, that unnecessary and inflationary expense would have represented 29 percent of all personal health care spending in that year.

Increased pressure on provider incomes comes not only from reduced reimbursements, but also from the administrative burden of dealing with many different insurance plans. With the era of cost containment also came greater competition between private insurers to offer employers acceptable benefits for their employees at the lowest price. One aspect of this competition is the creation of a broad and ever-changing menu of plans with different exclusions and different payment percentages for different health services. With this cacophony of payer fee schedules, the administrative overhead of providers in the U.S. has grown enormously, threatening to reduce provider incomes. (Woolhandler *et al.*, 2003, estimates administrative costs as 31 percent of provider revenue in the U.S. compared with 16 percent in Canada.) Providers have thus felt even more need to maintain their incomes through adaptation, and have consequently driven inflation in health care prices even further.

With the inclusion of provider adaptation in Figure 2 to explain health care inflation, a new loop is created: Loop R3. This loop describes the tug-of-war between payers restricting reimbursement in response to high health care costs, and providers adapting to these restrictions by effectively raising health care prices in an attempt to maintain their incomes. This loop has the effect of reducing the efficiency of health care spending and thus artificially raising the cost of health care to payers. The payers react to the magnified costs by seeking further restrictions on reimbursement, or by further denying insurance coverage. The net result is a reduction in health care assets and insurance coverage (through Loops B2 and B3, respectively), thus dampening growth in the extent of care. As shown below, this unintended chain of events might have been avoided or at least moderated had payers and providers not set Loop R3 in motion.

Baseline simulation and alternative tests of downstream behavior

In Figure 3 we present results from the baseline simulation for several of the model's key variables along with historical data. Results from the model are shown from 1960 through 2010.[4] We recognize that a couple of these data series are conceptually incomplete. In particular, the measure of health care investments does not include human capital, and the measure of the population at-risk fraction is based only on adults and on cardiovascular risk factors. Although more complete measures would likely show the same sorts of trends and have little or no effect on model findings, we would like to construct more complete data series, if possible, in future iterations of our model.

Having established the model's ability to do a good job of reproducing historical trends for a variety of key variables, let us examine how a few of the key feedback loops in Figure 2—in particular, those depicting the reactions of payers and providers—contribute to the overall simulated behavior. Shown in Figure 4 are results from the base run alongside results from alternative simulations for 1960–2010 in which one or more of these feedback loops has been cut. The assumptions and results for the simulations are as follows:

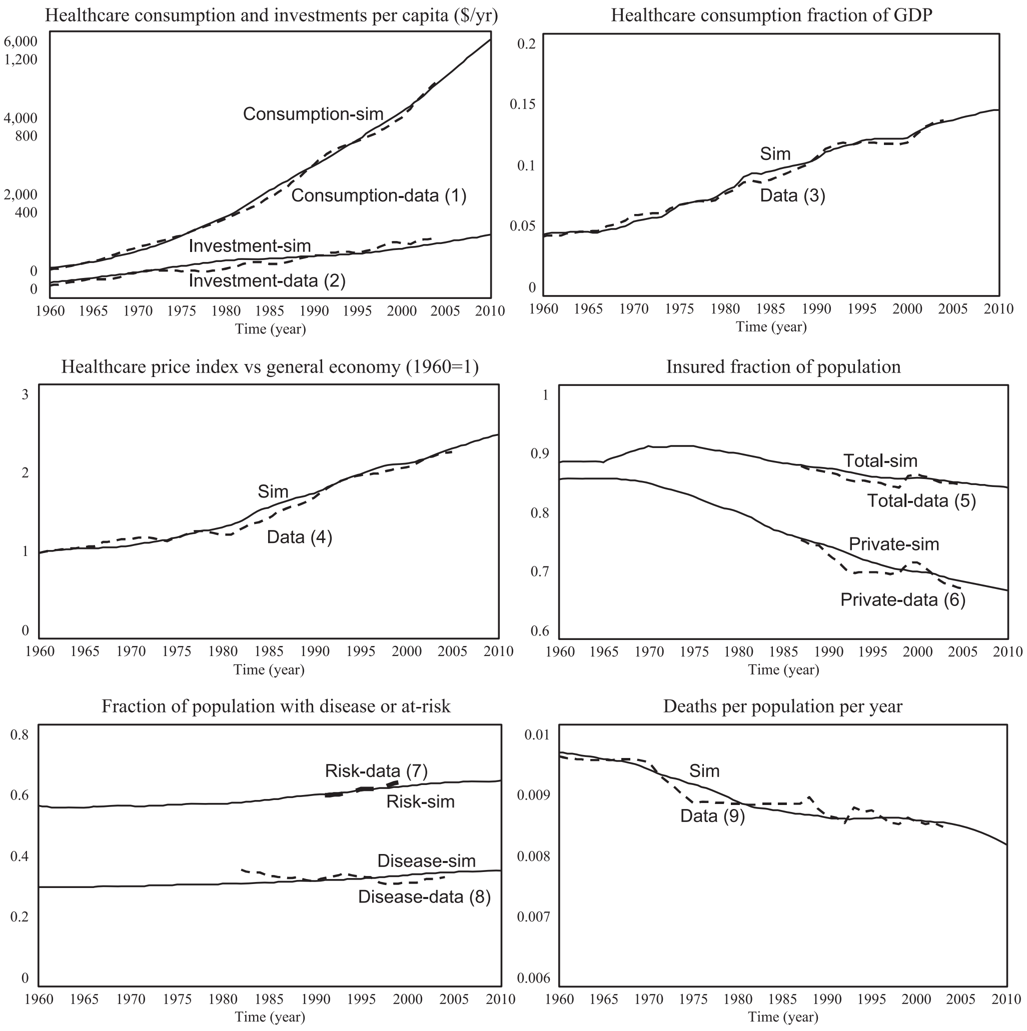


Fig. 3. Baseline simulation and historical data

Data sources are as follows:

(1) NHEA personal health care spending, 1960–2004 annual, divided by population and by GDP deflator (2000 = 1)
(2) NHEA investments in structures and equipment, 1960–2004 annual, divided by population and by GDP deflator
(3) NHEA personal health care spending divided by GDP, 1960–2004 annual
(4) BLS medical care CPI (1960 = 1) divided by general economy CPI (1960 = 1), 1960–2005 annual
(5) Census fraction of population of all ages covered by private or government health insurance, 1987–2005 annual
(6) Census fraction of population of all ages covered by private health insurance, 1987–2005 annual
(7) BRFSS fraction of adults who report having at least one of five specified cardiovascular risk factors, 1991–1999 odd years
(8) NHIS fraction of population of all ages who report their health as good, fair, or poor (i.e., not excellent or very good), 1982–2004 annual
(9) NVSR total deaths per year divided by population, 1960–1980 every 5 years, 1980–2003 annual

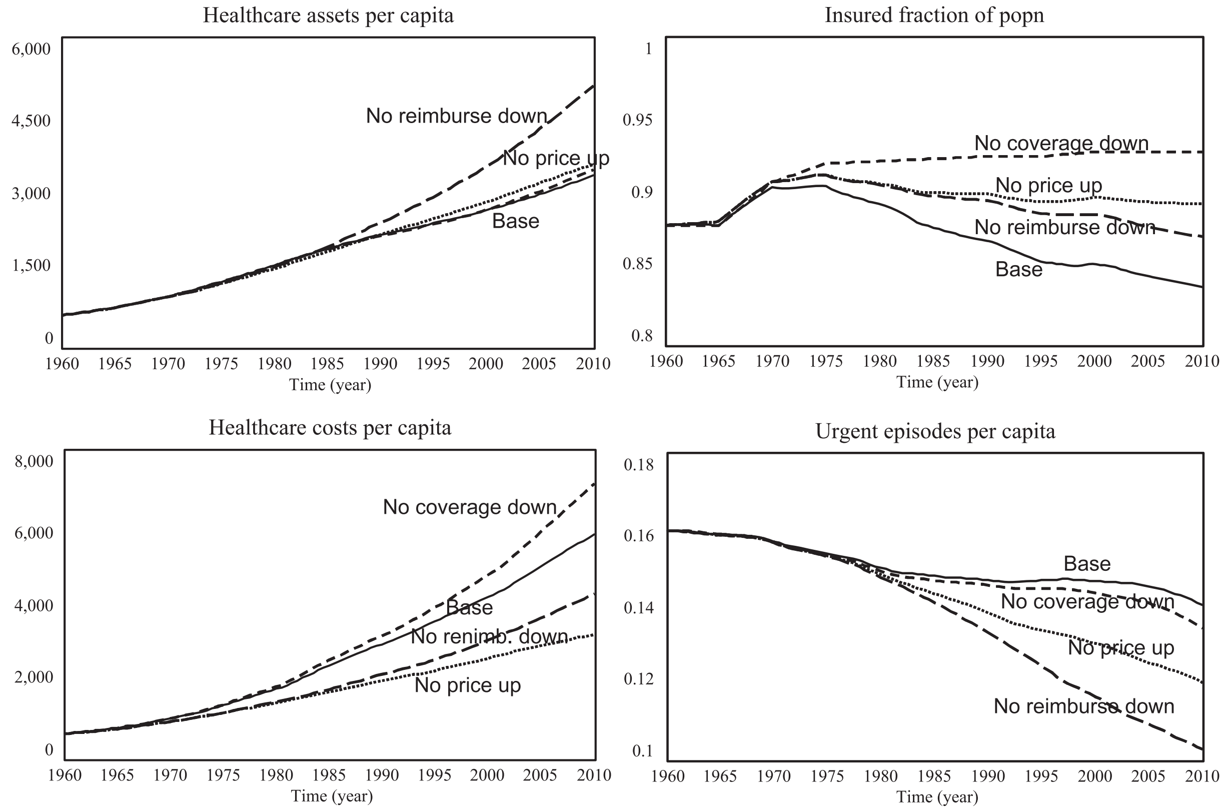


Fig. 4. Simulations exploring how reactions of payers, providers, and investors have determined health care system behavior

No coverage down. In this simulation, employers do not react to high health care costs by denying private insurance coverage; Loop B3 is cut. As a result, the insured fraction of the population does not decline, as it does in the base run; instead it continues to climb gradually to reflect the increasing availability of Medicaid coverage to those with lower incomes. With increased coverage, the extent of care—particularly disease management—is improved, and the rate of urgent episodes is therefore lower than in the base run. But this more extensive care costs more than it saves, and thus health care costs per capita increase relative to the base run. This outcome would seem to suggest that employers who have denied coverage to their employees have thereby saved money. Note, however, that the costs in the model do not include the sick days and losses of productivity that are much more likely to occur when disease is not well managed. This is why some of the nation's employers are taking another tack, providing free or low-cost primary health care in their own offices as a way of improving productivity and catching health problems before they get more serious and require expensive outside care (Freudenheim, 2007).

No price up. In this simulation, providers do not react to restrictions on reimbursement by raising their fees for a given quality of service; Loop R3 is cut. As a result, health care costs per capita grow much less than in the base run. Lower costs mean fewer revenues available for reinvestment but also less restriction of reimbursements and coverage. Because reimbursements are more stable, the investment rate does not decline as much as in the base run, and so, despite the decline in the revenue base, assets per capita increase no less than in the base run and even a bit more. With lower costs, there is also much less denial of insurance coverage. Because of the greater insurance coverage and the slightly greater assets, the extent of care is improved and urgent episodes per capita are reduced relative to the base run.

This simulation points to the importance of the “dysfunctional” Loop R3, but its results should not be taken too literally or as a prescription for policy. Legislators seeking to stabilize health care costs might be tempted to limit the autonomy of providers when it comes to billing and compensation, requiring that they be paid a fixed amount (as is done in some managed care organizations), perhaps through a single government payer (as is done in many countries). Some providers in the U.S. might welcome the predictability and reduced administrative burden such a simplified payment system would bring. Others, however, are likely to protest such loss of autonomy, especially the many who expect (and whose adaptive behavior to date has been based on the expectation of) high incomes in return for their long years of education and training. A national fixed-price policy might therefore be met by a decline in the supply of providers—an increased rate of retirement and decreased influx of medical students—leading perhaps to a severe shortage. In terms of Figure 2, if the adaptive responses of providers were no longer permitted, we might see a decline in the human capital component of health care assets; that is, a strengthening of Loop B2. Such a reaction could conceivably cause a fixed-price policy to do more harm than good, if the reaction were strong enough.

No reimburse down. In this simulation, employers and other payers do not react to high health care costs by restricting reimbursements; Loops B2 and R3 are cut. The stable reimbursement climate encourages more investment in assets as a fraction of revenues and also defuses the dysfunctional tug-of-war between payers and providers that leads to price inflation in Loop R3. The cutting of Loop R3 does keep health care costs down (as in the *No price up* simulation), but in so doing reduces health care revenues and therefore initially counteracts the effect of an increased investment fraction on asset formation, relative to the base run. By the 1980s, however, the stable investment fraction increasingly differentiates this scenario from the base run, and assets thus start to grow faster than in the base run. The rapid growth in assets per capita drives greater improvements in extent of care so that urgent episodes decline much more than in the base run. Also, the lower health care costs relative to the base run mean that there is less loss of insurance coverage, which improves the extent of care further.

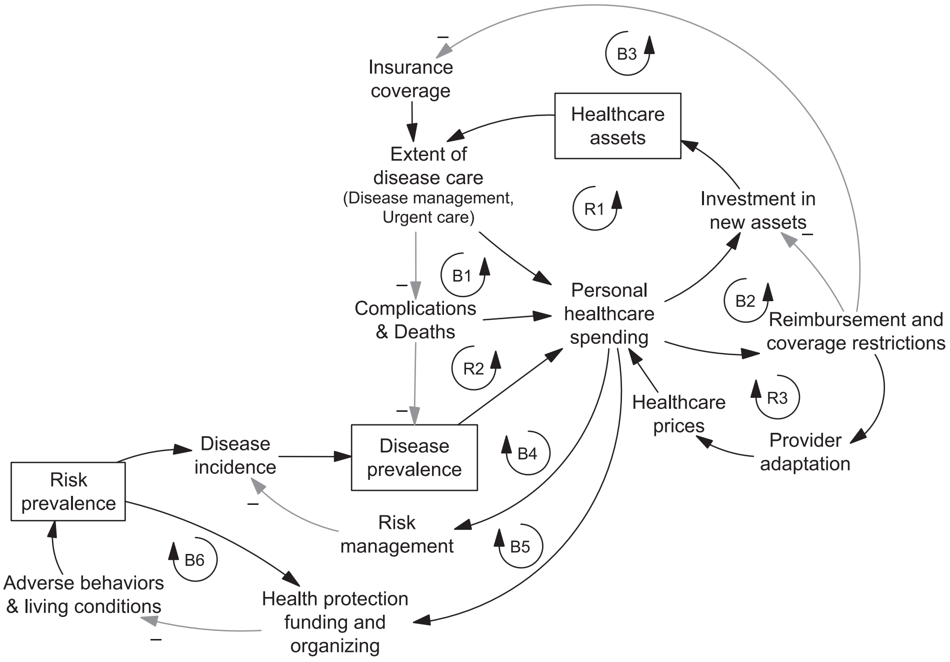
The *No reimburse down* simulation underscores the importance of the dysfunctional payer–provider interaction in Loop R3 and also points to the importance of the impact of payers on investors in Loop B2. But, as above, the results should not necessarily be viewed as having direct policy implications. They seem to suggest—perhaps counterintuitively—that health insurance should be stable and nonrestrictive in its reimbursements, so as to avoid behavioral backlashes that can trigger health care inflation and underinvestment. However, few policymakers in the U.S. would at present be willing to mandate that private payers must provide plans of only a certain sort, as such a mandate would be seen as interference in a matter of private choice. Perhaps, then, the mandate could apply only to the government's own insurance programs. (Government reimbursement practices are often copied by private insurers, and so with such an approach one may end up with the desired effect on the private sector without having to interfere with it.) Even so, many policymakers might fear that such a mandate would open the door not only to beneficial investments, but also to indiscriminate and wasteful ones, such as occurred most prominently before the era of cost containment. Still, it is interesting to consider whether a more generous and stable approach to reimbursement could not only combat illness better than the current restrictive approach, but do it more efficiently and perhaps even at lower cost.

The above analysis suggests that there are no easy downstream fixes to the problem of an under-performing and expensive health care system. It is one thing to understand the dysfunctional tug-of-war between payers and providers, but quite another to defuse it. We have addressed the lagging extent of care in our model by looking at the influences of health care assets and insurance coverage, but we have not explored improvements in the efficacy and safety of that care. Such improvements can include better information and decision-support systems, better payment incentives, and better clinical training. Local implementations of such improvements indicate their promise for reducing the burden of disease and providing more effective care for the health care dollar (IOM, 2001). One wonders, though, just how much we can hope to gain from such downstream measures, when they may appear to payers or providers as an even greater expense to bear (at least initially) and could therefore end up feeding into the system's divisiveness and dysfunction.

Potential upstream loops and tests of their behavior

Let us turn, then, to the upstream prevention of disease incidence, to see what promise it may hold for lessening our dependence on a costly and inefficient

Fig. 5. Extending the theory to allow for upstream responses


–
Insurance coverage
B3
Healthcare assets
Extent of disease care
(Disease management, Urgent care)
R1
Investment in new assets
–
B1
–
Complications & Deaths
Personal healthcare spending
B2
Reimbursement and coverage restrictions
R2
–
Healthcare prices
R3
Provider adaptation
Disease incidence
Disease prevalence
B4
Risk prevalence
–
Risk management
B5
B6
Adverse behaviors & living conditions
Health protection funding and organizing
–

Key to upstream feedback loops:
B4: Risk management proportional to downstream spending can help limit it
B5: Health protection proportional to downstream spending can help limit it
B6: Health protection (via sin taxes) proportional to risk prevalence can help limit it

system of downstream care (Fries *et al.*, 1998; McKinlay, 1979; McKinlay and Marceau, 2000). Illustrated in Figure 5 are two broad categories of such efforts: risk management for people already at risk, and health protection for the population at large. The literature identifies significant opportunities for medically oriented risk management for a variety of diseases, through improved nutrition and exercise, smoking cessation, and the appropriate use of drugs (Eyre *et al.*, 2004; Hajjar *et al.*, 2006; Leonhardt, 2007). For example, the fraction of people with hypertension whose condition is considered under control stood at 29 percent for 1999–2002, up only a few percentage points from 25 percent for 1988–1991 (Hajjar *et al.*, 2006).

The literature also describes opportunities for socially oriented health protection, which may include efforts to change adverse behaviors and mitigate unhealthy conditions in homes, schools, workplaces, and neighborhoods and to alter macroeconomic forces and the media so that they are more health promoting (Northridge *et al.*, 2003; Gerberding, 2005; Yach *et al.*, 2005; Simon, 2006; CDC, 1999, 2006; IOM, 2002; Smedley and Syme, 2000; Wilkinson and Marmot, 2003; Evans *et al.*, 1994; Hanna and Coussens, 2001). Note that, unlike downstream interventions, health protection efforts rely on the actions of individuals and organizations most of whom are not health care professionals.

What can the data tell us about the history of upstream spending? Much upstream work involves population-based public health efforts emphasizing health promotion and disease prevention. Public health spending has grown as a fraction of total NHEA spending from 1.5 percent in the early 1960s to a fairly constant 3 percent since the early 1990s. Another contributor to upstream spending is risk management. We have data from various reports, both public (e.g., NIHCM, 2002) and proprietary, that have allowed us to assemble a partial time series on spending on drugs for treating hypertension and high cholesterol. These data suggest that the use of these drugs grew from a negligible amount before 1980 to at least 1.5 percent of NHEA spending by 2004. Risk management in total would also include prescribed treatments for weight loss and smoking cessation, for which we have not yet assembled historical data. Thus, we estimate that upstream spending has grown to more than 4.5 percent (=3 percent population-based public health + more than 1.5 percent risk management) of total health spending. This amount is larger than the 3 percent upstream spending that was estimated in a 1991 report (Brown *et al.*, 1991). The data thus show that upstream spending has grown as a fraction of total health spending since 1960, even if it is still a relatively small fraction. This conclusion is significant because it stands counter to an impression we had before this study, that upstream spending had in recent decades been "squeezed out" by downstream spending (Homer and Hirsch, 2006).

Three balancing feedback loops have been included in Figure 5 and in our model to indicate how, in general terms, efforts in risk management and health protection might be funded or resourced more systematically and in proportion to indicators of capability or relative need. Funding is not the only prerequisite for such efforts, which also depend upon the enthusiasm and organization of the people involved (providers and patients in the case of risk management, and the general public in the case of health protection), but it is the leading requirement for most initiatives. Loop B4 suggests that funding for programs promoting risk management could be made proportional to spending on downstream care, so that when downstream care grows, funding for risk management would grow as well. Loop B5 suggests something similar for health protection, supposing that government budgets and philanthropic investments for health protection could be set in proportion to recent health care spending. Loop B6 takes a different approach to the funding of health protection, linking it not to health care spending but to risk prevalence (the stock which health protection most directly seeks to reduce). The linkage to risk prevalence can be made fiscally through "sin taxes" on unhealthy items, such as cigarettes (already taxed throughout the U.S. to varying extents; see Lindblom, 2006) and fatty foods (Marshall, 2000). In theory, the optimal magnitude of such taxes may be rather large in some cases, as the taxes can be used both to discourage unhealthy activities and promote healthier ones (O'Donoghue and Rabin, 2006).

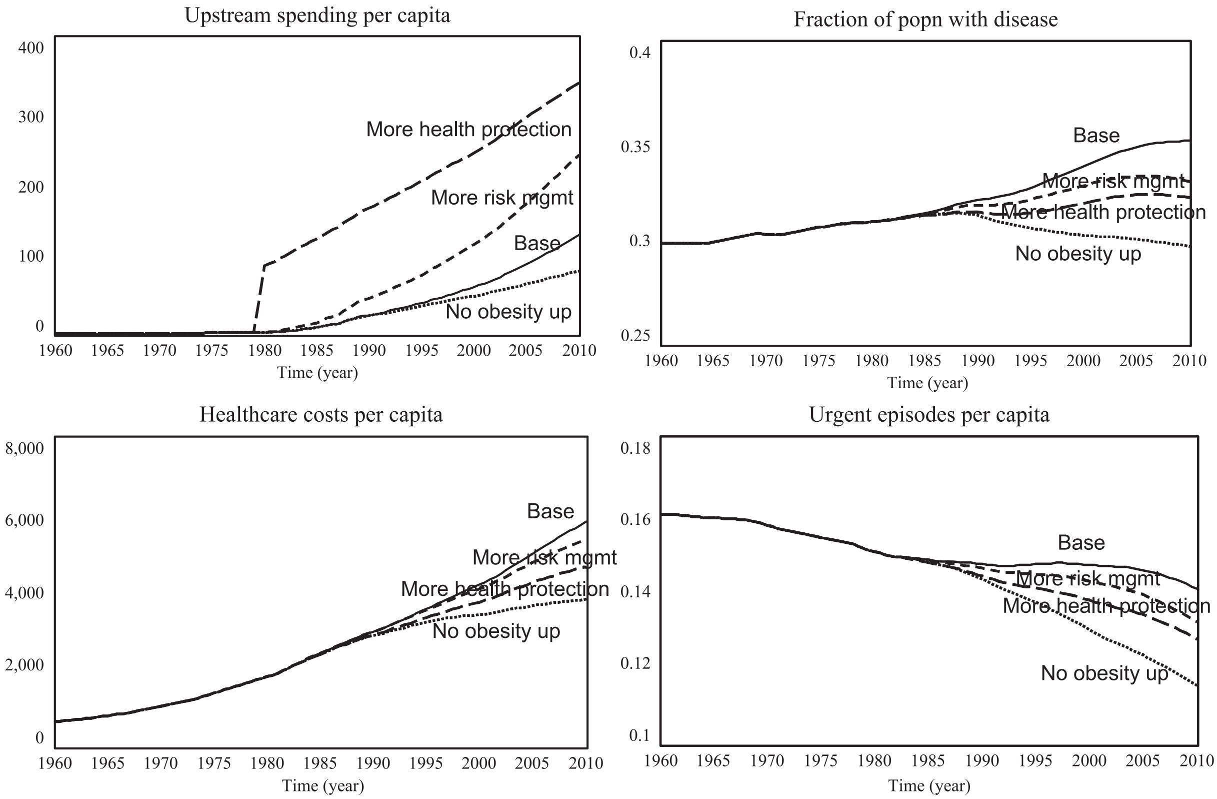


Fig. 6. Simulations exploring how upstream responses could alter health care system behavior

Presented in Figure 6 are results from simulations in which we ask how much the prevalence and burden of disease might have been diminished relative to the base run (through the year 2010) if greater upstream efforts at risk management or health protection had been made starting in 1980. These results may also be compared to a scenario, *No obesity up*, in which we assume that the base run's exogenous increase in the onset of risk by 35 percent from 1980 to 2005—representing a host of socioeconomic factors that have led to greater net caloric intake and obesity—had never occurred.

In these simulations we make assumptions about the degree to which upstream spending can affect rates of the onset of risk, reduction of risk, and onset of disease. In particular, we assume that maximum risk management could reduce the onset of disease by 40 percent and enhance reduction of risk by 40 percent, and that maximum health protection efforts could reduce onset of risk onset by 50 percent and enhance reduction of risk by 50 percent. For risk management, the assumptions, although uncertain, have been informed by studies focusing on the cost-effectiveness of risk management for patients with diabetes (CDC/DCEG, 2002; Hayashino *et al.*, 2004). For health protection, our assumptions are more uncertain, because relatively little is known about the required cost and potential impact of measures that could prevent the onset of various risk factors for disease; somewhat more is known in this regard about preventing smoking than about preventing obesity (Tengs *et al.*, 2001; Homer *et al.*, 2006). The Tengs analysis, focusing on a school-based antismoking program for young teens, estimated a cost of about $50 per student per year and projected long-term benefits in terms of reduced medical costs and increased quality-adjusted life years.

Because our assumptions about upstream efforts are associated with significant uncertainties, we do not purport here to provide accurate cost-effectiveness estimates, but only to illustrate how such estimates may be generated by our model. It is interesting to ask not only to what degree upstream efforts can improve health but also whether, and over what time frame, increases in upstream spending can be justified in terms of subsequent reductions in downstream spending. The model calculates upstream spending as the sum of risk management and health protection spending, and calculates downstream spending as the sum of all personal health care spending less spending on risk management. These measures of spending are accumulated over time, starting in 1980, as a way of quantifying overall costs and benefits; in the current model, no discount rate is applied to these cumulative measures.

No obesity up. This simulation is presented as a "best case" alternative history to the base run. Relative to the base run, the fraction of the population with disease grows more slowly during the 1980s, and this fraction declines from the 1990s onward rather than continuing to grow as in the base run. The result is much more progress starting in the late 1980s in reducing the number of urgent episodes, as well as a significant slowing in the growth of health care costs. This simulation indicates the extent to which increasing risk prevalence has undermined progress on health and has pushed health care costs upward since the late 1980s.

More risk management. In this simulation, the strength of the assumed linkage between personal health care spending (specifically, the non-urgent portion of that spending) and risk management is doubled relative to the base run; thus, the strength of Loop B4 in Figure 5 is doubled. By 2010, upstream spending per capita is increased by $108 relative to the base run (see Figure 6), and the effectiveness of risk management (in terms of reducing disease incidence and enhancing risk reduction) is increased to 51 percent of its assumed potential, versus 27 percent in the base run. The increase in risk management leads to slower growth of disease prevalence starting in the late 1980s. But, with the onset of risk left unaddressed in this scenario, disease prevalence does grow rather than decline. Urgent episodes and health care costs are somewhat improved relative to the base run, but not dramatically so. By 2010, a cumulative additional $359 billion in upstream spending since 1980 has led to a reduction in downstream spending of $1140 billion. The increased spending in risk management is not paid back immediately, however. Not until 1995, 15 years after the policy is initiated, does the cumulative reduction in downstream spending exceed the cumulative increase in upstream spending.

More health protection. In this simulation, health protection is much enhanced through a proportional funding program, starting in 1980, that devotes $5 to health protection for every $100 of personal health care spending; thus, Loop B5 in Figure 5 is activated. The result is an immediate $90 per capita increase in upstream spending in 1980, which grows to a $203 per capita increase by 2010 relative to the base run. By 2010, the effectiveness of health protection (in terms of reducing risk incidence and enhancing risk reduction) is increased from its baseline value of 19 percent to 48 percent of its assumed potential. The increase in health protection goes a long way but does not quite offset the adverse socioeconomic influences (such as changes in food and activity environments) that increase the onset of risk in the base run. As shown in Figure 6, this simulation produces improvements about halfway between the base run and *No obesity up* with regard to disease prevalence, urgent episodes, and personal health care costs. By 2010, a cumulative additional $1288 billion in upstream spending since 1980 has led to a reduction in downstream spending of $2750 billion. The breakeven year does not occur until 2002, however—22 years after the policy is initiated.

The 22-year payback period under *More health protection* is notably greater than the 15-year payback period in the *More risk management* simulation. Much of this additional payback time comes from the fact that health protection acts further upstream than risk management does, as seen in Figure 5. Some of the additional payback time for health protection is also likely a reflection of the fact that there is much earlier (1980s) upstream spending under the health protection scenario than there is in the risk management scenario. The 1980s spending is arguably less cost-effective than spending is during the 1990s, the period of most rapid growth in disease prevalence. As a partial test of this idea, we have performed another simulation in which the health protection program is implemented not in 1980 but in 1985. By 2010, a cumulative additional $1209 billion in upstream spending has led to a reduction in downstream spending of $2250 billion, with breakeven occurring in 2004, 19 years after the policy is initiated. Thus, the payback period is a few years less (19 versus 22) in this simulation, but the breakeven year has actually been pushed back further (2004 versus 2002).

In any event, whether the approach to upstream action is risk management or health protection, the model suggests that the payback time, purely in terms of health care costs, may be a relatively long one. It should be noted, however, that our model does not include losses in productivity to employers and society at large. Another SD model suggests that when these losses are taken into account, the payback on upstream action may shrink to a much shorter time period (Homer *et al.*, 2004), a length of time that may be acceptable to the public as well as to those employers in a position to put upstream efforts into effect.

A couple of broad conclusions may be drawn from the model simulations presented here in Figures 4 and 6. First, we see that cost-containment measures in the U.S. have thus far been futile, and they have done more to limit growth in the extent of care than they have to limit costs. In this sense, the existing market for health care services has been dysfunctional, and it would appear that societal measures to stabilize and simplify this market might be considered. Second, we see that progress is possible even within the current system to reduce costs and improve health through increased investments in upstream risk management and health protection measures. The financial payback on such investments may take some years but could ultimately be very large.

Conclusion

We have sought to explain why chronic illness is such a difficult problem to deal with and why the U.S. in particular has stumbled both in producing better health outcomes and in controlling the cost of illness.[5] Part of the problem comes from the growth in health risk that leads to greater incidence of disease, as exemplified by the rise in obesity. But another aspect of the problem is that progress in improving the treatment of existing illness seems to have stalled in recent years. Growth in health care assets has historically been a key driver of improved extent of care, but with rising costs reimbursement has become constrained, thereby creating uncertainty in the minds of potential investors regarding future revenues, and slowing investment in assets. Another driver of extent of care is insurance coverage, and private coverage, like reimbursement, has declined in response to rising costs. However, the impact of such decline has been mitigated by the availability of the government's Medicaid program as fallback coverage for many lower-income workers.

If rising costs are a great stumbling block to progress, why do they keep rising? Some of the increase is simply a reflection of past progress; namely, growth in health care assets, and increased longevity for those with chronic illness due to those health- and life-saving improvements. If this were the whole story, one could view the recent slowing in asset growth as part of an orderly process by which the health care system moves toward an acceptable maximum spending level relative to GDP. However, the data show that health care costs have continued to rise rapidly even as asset growth has slowed. The explanation for this continued rapid growth in costs may lie in income-maintaining adaptations by providers, who have been able to raise prices and service volumes for a given quality of care, especially in the less well-regulated outpatient sector. These adaptations have come in direct response to the attempts by payers to control costs through restrictions of reimbursement.

This tug-of-war between payers and providers, permitted by the current system of payment in the U.S., has had damaging effects in terms of limiting the supply of care, and it has also increased the administrative overhead of providers. In the absence of effective controls, health care costs as a fraction of GDP in the U.S. accelerated ahead of those in other industrialized countries starting in the 1980s, without delivering better care (Docteur and Oxley, 2003). In this sense, the entrepreneurial U.S. health care system which made such great progress in the past has now become bloated and inefficient. As many have come to realize, the time is overdue for a fundamental change in this dysfunctional system.

The difficulty of controlling costs and improving outcomes in the U.S. suggests the need for an innovative approach to health reform—one that emphasizes upstream efforts to reduce the health risks that may lead to chronic illness. While spending on population-based health protection and risk management programs has grown somewhat, it still represents a small fraction of total U.S. health care spending. Our model suggests that policies that shift the balance toward more upstream programs can have beneficial impacts on both health care costs and the population's health status. Although such upstream investments may take several years or even decades to come to fruition, it is important to recognize that improved health is a chief aspiration of all people and therefore deserves a commitment to strategies that will benefit both current and future generations.

Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

Notes

1. Despite the clear importance of cost controls in the evolution of U.S. health care since the 1970s, we have no hard data on how they have changed over time. Consequently, although reimbursement and coverage restrictions play a central role in Figure 2, they are not included explicitly in the simulation model. Instead, in the model we focus on *personal health care costs as a fraction of GDP* as the key factor to which other variables in Figure 2 react, including the reinvestment rate in new assets, private insurance coverage, and inflation in health care prices. Each one of these responses to increased

costs is modeled with a delay of 3 or 4 years, reflecting the adjustment times of the relevant stakeholders, who (depending upon the variable in question) may be employers, prospective investors, or health care providers.

2. We have modeled the fraction covered by any insurance (“total”) as the sum of (1) the fraction covered by private insurance and (2) the fraction covered by government insurance but not by private insurance. (A large fraction of the elderly have both Medicare and private insurance to supplement Medicare’s copay requirements and gaps in coverage. Because this population has both public and private insurance, one cannot simply add up the different categories of insurance to get the total covered population.) In line with the “fallback” argument in the main text, the government-only coverage is modeled as a fraction of those not covered by private insurance, a fraction specified by an exogenous time series. Based on Census data from 1987–2005, we estimate that this fallback fraction has risen only slightly in recent years, from 48 percent in the late 1980s to 50 percent by 2000. For the years preceding 1987, we have examined the NHEA, which provides a breakdown of health care spending by major payer category, to see how the balance of private and government-paid spending has changed over time. Based on these data in conjunction with the Census data for the late 1980s, we estimate that the fallback fraction sat at about 18 percent during the years preceding the establishment of Medicare and Medicaid in 1965 but rose rapidly thereafter to 37 percent in 1970 and 45 percent in 1975.
3. Although Loops B2 and B3 appear to act similarly, their impacts on disease management or urgent care, as subsets of disease care, are rather different. With regard to Loop B3, we have noted previously that the uninsured poor have greater access to urgent care, through hospital emergency departments, than they do to disease management. Thus, when a person loses insurance coverage, this will tend to lead to more of a reduction in disease management than in urgent care. Lacking disease management, this person becomes more prone to complications of disease leading to expensive hospitalization, the costs of which are shifted from the employer to hospitals and the general public. With regard to Loop B2, the trend in reduced investment appears to have been more benign, more genuinely cost saving due to limitations on urgent care. Our model-based analysis of NHEA data suggests that the fraction of investments directed to disease management rather than urgent care was roughly 30 percent through 1980 but by 2004 had increased to 35 percent. That is, investors appear to have moved gradually more in the direction of disease management and away from urgent care.
4. Simulation beyond 2005 requires assumptions for several different input time series. For future deaths and total population, we have used U.S. Census projections. For future growth in real GDP per capita, we have assumed a flat 1.9 percent per year, which was the average growth rate for 1985–2005. For future changes in risk onset and disease onset due to social

and economic influences beyond the scope of the model, we have assumed no further increase after 2005. For the government coverage fraction of people not covered by private insurance plans, we have assumed no further change beyond 2000, at which time we estimate the fraction at 50 percent. And for risk management spending as a fraction of total disease and risk management spending, we assume linear growth extrapolating from the past: from 1.5 percent in 1989 to 3.1 percent in 2004 to 3.8 percent in 2010.

5. The world's developing countries lack the resources to duplicate the expensive patterns of care that emerged in the U.S. and will need to find their own path. One study indicates that developing countries will require very different prevention strategies for cardiovascular diseases than those of higher income countries (Reddy and Yusuf, 1998). The authors suggest that currently low levels of cardiovascular risk factors in the large rural segments of the developing countries offer a window of opportunity for early and effective control of the epidemic. They state: "At the present levels of these risk factors in the developing countries, the approach would be predominantly non-pharmacological, population based, and lifestyle linked. This would largely avoid the biologic and economic costs of a pharmacological approach warranted by high levels of these risk factors in the developed countries."

References

ADA (American Diabetes Association). 2003. Economic costs of diabetes in the U.S. in 2002. *Diabetes Care* **26**(3): 917–932.

Brown R, Elixhauser A, Corea J, Luce B, Sheingod S. 1991. National expenditures for health promotion and disease prevention activities in the United States. *Report BHARC-013/91-019*. Battelle Medical Technology Assessment and Policy Research Center: Washington, DC.

BLS (Bureau of Labor Statistics). 2007. Consumer Price Index: all urban consumers (Series IDs CUUR0000SA0 and CUUR0000SAM). http://data.bls.gov/cgi-bin/surveymost?cu [24 January 2007].

CDC (Centers for Disease Control and Prevention). 1999. Ten great public health achievements: United States, 1900–1999. *Morbidity and Mortality Weekly Report* **48**(12): 241–243.

——. 2006. Health protection goals. http://www.cdc.gov/about/goals/goals.htm [5 October 2006].

CDC/BRFSS (Centers for Disease Control and Prevention). 2007. Behavioral Risk Factor Surveillance System. Prevalence data: general health. http://apps.nccd.cdc.gov/brfss/list.asp?cat=HS&yr=2005&qkey=1100&state=All [24 January 2007].

CDC/DCEG (CDC Diabetes Cost-Effectiveness Group). 2002. Cost-effectiveness of intensive glycemic control, intensified hypertension control, and serum cholesterol level reduction for Type 2 diabetes. *Journal of the American Medical Association* **287**(19): 2542–2551.

CDC/NCCDPHP (Centers for Disease Control and Prevention, National Center for

Chronic Disease Prevention and Health Promotion). 2007a. Chronic Disease Overview. http://www.cdc.gov/nccdphp overview.htm [24 July 2007].

——. 2007b. Overweight and Obesity: Obesity Trends. http://www.cdc.gov/nccdphp/dnpa/obesity/trend/index.htm [24 July 2007].

CDC/NHIS (Centers for Disease Control and Prevention). 2007. National Health Interview Survey, Series 10 (annual). Respondent-Assessed Health Status and Reported Chronic Conditions. http://www.cdc.gov/nchs/products/pubs/pubd/series/sr10/ser10.htm [24 January 2007].

CDC/NVSR (Centers for Disease Control and Prevention). 2006. *National Vital Statistics Reports*, Vol. 54(13). http://www.cdc.gov/nchs/data/nvsr/nvsr54/nvsr54_13.pdf. [24 January 2007].

CMS (Center for Medicare and Medicaid Services). 2007. National Health Expenditure Data. http://www.cms.hhs.gov/NationalHealthExpendData/02_ [24 January 2007].

Cutler DM, Rosen AB, Vijan S. 2006. The value of medical spending the United States, 1960–2000. *New England Journal of Medicine* **355**(9): 920–927.

Docteur E, Oxley H. 2003. Health-care systems: lessons from the reform experience. Organization for Economic Cooperation and Development (OECD), Economics Department Working Papers No. 374. http://www.oecd.org/eco [24 July 2007].

Dominick KL, Ahern FM, Gold CH, Heller DA. 2002. Relationship of health-related quality of life to health care utilization and mortality among older adults. *Aging Clinical and Experimental Research* **14**: 499–508.

Eckholm E and *New York Times* staff. 1993. *Solving America's Health-Care Crisis.* Times Books: New York.

Evans RG, Barer ML, Marmor TR. 1994. *Why are Some People Healthy and Others Not? The Determinants of Health of Populations.* A de Gruyter: New York.

Eyre H, Kahn R, Robertson RM *et al.* 2004. Preventing cancer, cardiovascular disease, and diabetes: a common agenda for the American Cancer Society, the American Diabetes Association, and the American Heart Association. *Circulation* **109**: 3244–3255.

Freudenheim M. 2007. Company clinics cut health costs. *New York Times* 14 January.

Fries JF, Koop CE, Sokolov J, Beadle CE, Daniel W. 1998. Beyond health promotion: reducing need and demand for medical care. *Health Affairs* **17**(2): 70–84.

Gerberding JL. 2005. Protecting health: the new research imperative. *Journal of the American Medical Association* **294**(11): 1403–1406.

Greenlund KJ, Zheng ZJ, Keenan NL *et al.* 2004. Trends in self-reported multiple cardiovascular disease risk factors among adults in the United States, 1991–1999. *Archives of Internal Medicine* **164**: 181–188.

Hajjar I, Kotchen JM, Kotchen TA. 2006. Hypertension: trends in prevalence, incidence, and control. *Annual Review of Public Health* **27**: 465–490.

Hanna KE, Coussens C. 2001. *Rebuilding the Unity of Health and the Environment: A New Vision of Environmental Health for the 21st Century.* Institute of Medicine, National Academy Press: Washington, DC.

Hayashino Y, Nagata-Kobayashi S, Morimoto T *et al.* 2004. Cost-effectiveness of screening for coronary artery disease in asymptomatic patients with Type 2 diabetes and additional atherogenic risk factors. *Journal of General Internal Medicine* **19**(12): 1181–1191.

Heirich M. 1999. *Rethinking Health Care.* Westview Press: Boulder, CO.

Hirsch G, Homer J, McDonnell G, Milstein B. 2005. Achieving health care reform in the United States: toward a whole-system understanding. In *23rd International Conference of the System Dynamics Society*, Boston, MA.

Hirsch GB, Bergan TA, Goodman MR. 1975. Examining Alternatives for Improving the Nation's Oral Health: A System Dynamics Model of the Dental Care Delivery System. A report to the Division of Dentistry, Bureau of Health Manpower, DHEW: Washington, DC.

Hirsch GB, Immediato CS. 1999. Microworlds and generic structures as resources for integrating care and improving health. *System Dynamics Review* **15**(3): 315–330.

Hirsch GB, Myers R. 1975. Designing Strategies for Particular Health Problems: The Indiana Cardiovascular Disease Model. Indiana Health Planning and Development Agency.

Hoffman C, Rice D, Sung H. 1996. Persons with chronic conditions: their prevalence and costs. *Journal of the American Medical Association* **276**(18): 1473–1479.

Homer J, Hirsch G, Minniti M, Pierson M. 2004. Models for collaboration: how system dynamics helped a community organize cost-effective care for chronic illness. *System Dynamics Review* **20**(3): 199–222.

Homer JB. 1996. Why we iterate: scientific modeling in theory and practice. *System Dynamics Review* **12**(1): 1–19.

Homer JB, Hirsch GB. 2006. System dynamics modeling for public health: background and opportunities. *American Journal of Public Health* **95**: 452–458.

Homer JB, Milstein B. 2002. Communities with multiple afflictions: a system dynamics approach to the study and prevention of syndemics. In *Proceedings of the 20th International Conference of the System Dynamics Society*, Palermo, Italy. System Dynamics Society. http://www.systemdynamics.org/conferences/2002/papers/Homer1.pdf [6 August 2007].

Homer J, Milstein B. 2004. Optimal decision making in a dynamic model of community health. In *37th Hawaii International Conference on System Sciences*. IEEE: Waikoloa, HI. http://csdl.computer.org/comp/proceedings/hicss/2004/2056/03/205630085a.pdf [24 July 2007].

Homer J, Milstein B, Dietz W, Buchner D, Majestic E. 2006. Obesity population dynamics: exploring historical growth and plausible futures in the U.S. In *24th International Conference on SD Society*, Nijmegen, Netherlands.

IOM (Institute of Medicine). 2001. *Crossing the Quality Chasm: A New Health System for the 21st Century*. National Academy Press: Washington, DC.

——. 2002. *The Future of the Public's Health in the 21st Century*. National Academy Press: Washington, DC.

Jones AP, Homer JB, Murphy DL, Essien JDK, Milstein B, Seville DA. 2006. Understanding diabetes population dynamics through simulation modeling and experimentation. *American Journal of Public Health* **96**(3): 488–494.

Leonhardt D. 2007. What's a pound of prevention really worth? *New York Times* 24 January.

Lindblom E. 2006. State cigarette excise tax rates and rankings. Campaign for Tobacco-Free Kids: Washington, DC. http://www.tobaccofreekids.org/research/factsheets/pdf/0097.pdf [5 February 2007].

Luginbuhl WH, Forsyth BR, Hirsch GB, Goodman MR. 1981. Prevention and rehabilitation as a means of cost containment: the example of myocardial infarction. *Journal of Public Health Policy* **2**(2): 103–115.

Mackenbach JP. 1994. The epidemiologic transition theory. *Journal of Epidemiology and Community Health* **48**(4): 329–331.

Marshall T. 2000. Exploring a fiscal food policy: the case of diet and ischaemic heart disease. *British Medical Journal* **320**: 301–305.

Mathers CD, Loncar D. 2006. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Medicine* **3**(11): 2011–2030.

McKinlay JB. 1979. A case for refocusing upstream: the political economy of illness. In *Patients, Physicians, and Illness: A Sourcebook in Behavioral Science and Health* (3rd edn.), Jaco EG (ed.). Free Press: New York; 9–25.

McKinlay JB, Marceau LD. 2000. Upstream healthy public policy: lessons from the battle of tobacco. *International Journal of Health Services* **30**(1): 49–69.

NIHCM (National Institute for Health Care Management Foundation). 2002. Prescription Drug Expenditures in 2001: Another Year of Escalating Costs. Washington, DC. http://www.nihcm.org/spending2001.pdf [24 January 2007].

NIH/NHLBI (National Institutes of Health, National Heart, Lung and Blood Institute). 2004. Morbidity and Mortality: 2004 Chart Book on Cardiovascular, Lung, and Blood Diseases. Washington, DC. http://www.nhlbi.nih.gov/resources/docs/cht-book.htm [24 July 2007].

Northridge ME, Sclar ED, Biswas P. 2003. Sorting out the connections between the built environment and health: a conceptual framework for navigating pathways and planning healthy cities. *Journal of Urban Health* **80**(4): 556–568.

NSC (National Safety Council). 2006. What are the odds of dying? (Odds of death due to injury). http://www.nsc.org/lrs/statinfo/odds.htm [24 January 2007].

O'Donoghue T, Rabin M. 2006. Optimal sin taxes. *Journal of Public Economics* **90**: 1825–1849.

Olshansky SJ, Ault AB. 1986. The fourth stage of the epidemiologic transition: the age of delayed degenerative diseases. *Milbank Quarterly* **64**(3): 355–391.

Ratanawijitrasin S. 1993. The dynamics of health care finance: a feedback view of system behavior. PhD dissertation, State University of New York at Albany, NY.

Reddy KS, Yusuf S. 1998. Emerging epidemic of cardiovascular disease in developing countries. *Circulation* **97**: 596–601.

Simon E. 2006. Bosses push staff to eat right, exercise. *Associated Press (New York)* 4 December.

Smedley BD, Syme SL. 2000. *Promoting Health: Intervention Strategies from Social and Behavioral Research*. Institute of Medicine. National Academy Press: Washington, DC.

Starr P. 1982. *The Social Transformation of American Medicine*. Basic Books: New York.

Tengs TO, Osgood ND, Chen LL. 2001. The cost-effectiveness of intensive national school-based anti-tobacco education: results from the tobacco policy model. *Preventive Medicine* **33**: 558–570.

Thorpe KE, Florence CS, Howard DH, Joski P. 2005. The rising prevalence of treated disease: effects on private health insurance spending. *Health Affairs (Web Exclusive)* **W5**: 317–325.

U.S. Census Bureau. 2007. Health Insurance Coverage Status and Type of Coverage by Age: 1987 to 2005 (Table HI-7). http://www.census.gov/hhes/www/hlthins/historic/hihistt7.html [24 January 2007].

Weisbrod BA. 1991. The health care quadrilemma: an essay on technological change, insurance, quality of care, and cost containment. *Journal of Economic Literature* **29**: 523–552.

WHO (World Health Organization). 2002. Innovative Care for Chronic Conditions: Building Blocks for Action. http://www.who.int/chp/knowledge/publications/icccglobalreport.pdf [24 July 2007].

——. 2005. Preventing Chronic Diseases: A Vital Investment. http://www.who.int/chp/chronic_disease_report/en [24 July 2007].

Wild S, Roglic G, Green A, Sicree R, King H. 2004. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care* **27**: 1047–1053.

Wilkinson RG, Marmot MG (eds). 2003. *The Solid Facts: Social Determinants of Health* (2nd edn.). Centre for Urban Health. World Health Organization: Copenhagen.

Woolhandler S, Campbell T, Himmelstein DU. 2003. Costs of health care administration in the United States and Canada. *New England Journal of Medicine* **349**: 768–775.

Yach D, Hawkes C, Gould CL, Hoffman KJ. 2004. The global burden of chronic diseases: overcoming impediments to prevention and control. *Journal of the American Medical Association* **291**(21): 2616–2622.

Yach D, McKee M, Lopez AD, Novotny T. 2005. Improving diet and physical activity: 12 lessons from controlling tobacco smoking. *British Medical Journal* **330**: 898–900.